A Novel Pansharpening Framework Based on the Atangana-Baleanu Fractal-Fractional Derivative

Husam Yahya[1], Nadia M. G. Al-Saidi[2], Arkan J. Mohammed[3], and Yeliz Karaca[4(✉)]

[1] Ministry of Higher Education and Scientific Research, Baghdad, Iraq
Husam.yahya@mohesr.edu.iq
[2] College of Applied Sciences, University of Technology, Baghdad, Iraq
nadia.m.ghanim@uotechnology.edu.iq
[3] Department of Mathematics, College of Science, Al-Mustansiriah University, Baghdad, Iraq
drarkanjasim@uomustansiriyah.edu.iq
[4] Department of Mathematics and Department of Neurology, University of Massachusetts Chan Medical School (UMASS), 55 Lake Avenue North, Worcester, MA 01655, USA
yeliz.karaca@ieee.org

Abstract. Combining spectral and spatial information from multiple sensors into a single, analysis-ready product is among the most persistent challenges in remote sensing and medical imaging. Pansharpening—the fusion of a high-resolution panchromatic band with lower-resolution multispectral imagery—is particularly demanding: improving spatial detail tends to compromise spectral accuracy, and preserving spectral fidelity tends to blur spatial structure. Classical methods based on component substitution, multiresolution analysis, or modulation transfer rely on integer-order operators whose local support cannot capture the non-local memory effects and scale-invariant structures present in natural terrain. We introduce ABR-FFD, a pansharpening framework that replaces the conventional Gaussian smoothing kernel with a Mittag-Leffler low-pass filter derived from the Atangana-Baleanu fractal-fractional derivative. The kernel is evaluated via exact Mittag-Leffler series summation without asymptotic simplification, and its two parameters—the fractional order α and fractal dimension τ—are jointly optimised per scene through a 15×15 grid search over a fixed 7×7 neighbourhood. A multiplicative injection rule augmented with local mean correction and a two-scale coarse-detail term preserves inter-band spectral ratios at every pixel. Evaluated on twenty WorldView-3 scenes from the PanCollection benchmark using Wald's reduced-resolution protocol against seven classical comparison methods, ABR-FFD simultaneously achieves the lowest ERGAS (4.768), the best spectral angle (SAM = 5.798°, winning on all twenty individual scenes), and the highest SSIM (0.8761). These three spectral-fidelity metrics measure fundamentally different aspects of fusion quality and have not previously been optimised simultaneously by

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16758, pp. 317–343, 2027.
https://doi.org/10.1007/978-3-032-30497-1_20

a single classical method. An ablation study identifies the local mean correction as the sole driver of the spectral angle gain, contributing a reduction of 0.036° relative to the uncorrected baseline, while neighbourhood-size analysis establishes a non-monotone relationship between kernel support and ERGAS, with 7×7 as the unique optimum for the WV3 4: 1 resolution ratio. These findings establish fractal-fractional operators as a principled and competitive alternative to wavelet-based smoothing in satellite image fusion.

Keywords: Pansharpening · Fractal-fractional calculus · Atangana-Baleanu derivative · Mittag-Leffler kernel · Remote sensing · Image fusion · WorldView-3

1 Introduction

Earth observation satellites now routinely acquire the same scene at different spatial and spectral resolutions, and turning those separate acquisitions into a single analysis-ready product has become one of remote sensing's central practical problems [1,3]. Pansharpening is the most common instance: a high-resolution panchromatic band is merged with lower-resolution multispectral bands to yield imagery that is simultaneously geometrically sharp and spectrally faithful [2,4]. For sensors such as WorldView-3 and GaoFen-2, where the resolution gap reaches 4:1, striking an acceptable spatial-spectral balance becomes considerably more difficult.

Classical pansharpening methods fall into three broad families, each with well-documented limitations. In component substitution (CS)—IHS [5], PCA [6], and Gram–Schmidt (GS) [7]—the MS bands are projected into a decorrelated space, one component is replaced by the PAN image, and the inverse transform is applied. The spatial sharpness achieved this way comes at a cost: since the replaced component rarely matches the PAN spectral response, the reconstructed bands typically carry systematic colour distortions [4]. Multiresolution analysis (MRA) methods such as AWLP [8] and HPF [9] take a different route altogether, extracting PAN detail at multiple scales and adding it back into the MS bands directly; spectral preservation is generally better, though fine-textured regions remain problematic when the filter does not align well with local edge orientations [12]. Modulation-based methods, exemplified by SFIM [10], scale each MS band by the local PAN-to-smooth-PAN ratio—a conceptually elegant idea that produces a reasonable spatial-spectral compromise. The bottleneck, however, is the smoother: a fixed Gaussian kernel responds identically at edges and in flat regions, so performance saturates regardless of how carefully the bandwidth is tuned [11]. CNN, GAN, and transformer architectures [13–15] have pushed benchmark scores considerably higher, yet each requires a large labelled training corpus, carries non-trivial inference cost, and may generalise poorly when the target sensor or scene type is not well represented in training.

Operators such as Sobel and Canny inherit this strictly local character, which makes them computationally attractive but spatially myopic: high-frequency

noise is amplified while gradual, long-range variations are underweighted [16,17]. In fusion applications the result is familiar—staircase artefacts and loss of structural coherence—because each output pixel sees only its immediate neighbours and has no access to the surrounding textural context.

Fractional calculus extends differentiation to non-integer orders $\alpha \in \mathbb{R}$, and the resulting operators carry a property that is genuinely useful here: they integrate function values over an extended interval rather than at a point, giving each output location implicit access to a wider spatial context and making them inherently more sensitive to texture than their integer-order counterparts [2,18–20]. The classical Riemann–Liouville and Caputo formulations are not without problems, though. Both employ power-law kernels that diverge as $\xi \to t$, and near sharp image boundaries this singularity can trigger numerical instabilities that reintroduce the very staircase artefacts one was hoping to avoid [21]. Atangana and Baleanu's 2016 derivative replaces that power-law with a Mittag-Leffler kernel, which stays bounded everywhere while preserving the non-local integration range [21,22]:

$$\frac{\partial I}{\partial x} = \lim_{h \to 0} \frac{I(x+h, y) - I(x, y)}{h}. \tag{1}$$

Fractal geometry, formalised by Mandelbrot [23], offers a complementary perspective: it quantifies self-similar, scale-invariant structure through the fractal dimension D, capturing geometric complexity that ordinary Euclidean measures miss. Such properties are ubiquitous in natural and man-made terrain as seen from orbit. Within image analysis, fractal descriptors have found use in texture classification, segmentation, compression, medical imaging, and remote sensing interpretation; in remote sensing specifically they have been applied to land-cover characterisation, river network and coastline analysis, image-type classification, and hyperspectral dimensionality reduction [24–26]. Adding D as an operator parameter allows the filter to remain sensitive to irregular, self-similar texture rather than smoothing it away as purely Euclidean operators tend to do. Pairing the AB derivative with a fractal dimension parameter τ therefore yields a two-parameter operator that simultaneously accounts for non-local memory and self-similar terrain geometry—a combination well suited to the multi-scale, heterogeneous content typical of satellite imagery. Against this background, we propose ABR-FFD, a pansharpening framework that substitutes the Gaussian smoothing kernel in SFIM with a Mittag-Leffler low-pass filter derived from the Atangana-Baleanu fractal-fractional derivative, evaluated via exact series summation without asymptotic approximation. The two parameters (α, τ) are jointly optimised per scene through a 15×15 grid search, and the pipeline incorporates a local mean correction and two-scale coarse-detail injection to improve spectral angle preservation. This paper makes four contributions: (i) the derivation of an isotropic Mittag-Leffler low-pass kernel with exact series coefficients, free of asymptotic substitution, for spatial smoothing in the SFIM modulation framework; (ii) a two-parameter (α, τ) optimisation strategy over a dense grid, with the neighbourhood size fixed at the experimentally confirmed optimum of 7×7; (iii) an evaluation on twenty WorldView-3 scenes from the PanCol-

lection benchmark [15] against seven classical methods spanning all three pansharpening families, establishing best performance on ERGAS, SAM, and SSIM simultaneously; and (iv) an ablation study and neighbourhood-size analysis that quantify the individual contribution of each design element.

The remainder of this paper is organised as follows. Section 2 surveys related work in pansharpening and fractional image processing. Section 3 presents the mathematical background. Section 4 derives the proposed ABR-FFD operator and fusion pipeline. Section 5 reports experimental results, and Sect. 6 discusses findings and limitations. Finally, Sect. 7 concludes the paper.

2 Literature Review

This section surveys the principal families of pansharpening methods and recent advances in fractional and fractal-based image processing.

2.1 Classical Pansharpening Methods

CS methods represent the earliest class of pansharpening algorithms and have endured largely because of their computational efficiency. Carper et al. [5] proposed IHS fusion: the MS bands are converted to intensity-hue-saturation space, the intensity channel is replaced by the histogram-matched PAN image, and the inverse transform is applied; spatial sharpness is high, but the spectral distortion introduced scales with the mismatch between the PAN response and the intensity channel. Chavez et al. [6] used principal component analysis in a similar vein, substituting the first component with PAN and inheriting the same spectral bias. Laben and Brower's Gram–Schmidt (GS) approach [7] reduces this bias somewhat by first synthesising a simulated PAN from the MS bands before substitution, though the improvement is incomplete for spectrally heterogeneous scenes.

MRA methods take a fundamentally different approach, injecting spatially decomposed PAN detail directly into the MS bands. Schowengerdt [9] introduced high-pass filtering (HPF) as an early example; Otazu et al. [8] refined it through the à trous wavelet-based AWLP, which adapts injection gains per band using spectral response functions, generally yielding good spectral preservation. One known weakness of AWLP is that the wavelet basis can be poorly matched to dominant edge orientations in fine-textured regions, leading to visible ringing artefacts in some scenes. Liu's SFIM [10] occupies the boundary between MRA and modulation approaches: each MS band is rescaled by the ratio of PAN to a low-pass approximation of PAN, which is straightforward to implement and tends to produce acceptable spectral fidelity. The persistent problem is that the Gaussian used for the low-pass step cannot adapt to local image structure, and SFIM's typically modest SCC scores reflect this rigidity directly.

2.2 Deep Learning and Model-Based Approaches

Deep learning approaches—including CNNs, GANs, and more recently transformer-based networks—have delivered substantial gains on standard pansharpening benchmarks [13–15]. The practical limitations are nonetheless real: collecting sufficient labelled training data is non-trivial, inference times scale with model size, and out-of-distribution generalisation (different sensor, different biome) can be unpredictable. These constraints have kept interest alive in classical model-based approaches, whose assumptions are explicit and whose failure modes can often be anticipated from the mathematical structure alone.

2.3 Fractional Calculus in Image Processing

Fractional-order operators have been an active topic in image processing for roughly a decade. Riemann–Liouville and Grünwald–Letnikov derivatives have been used for edge detection in remote sensing, where capturing multi-scale responses within a single operator is particularly valuable [16]. Pu et al. [20] demonstrated that fractional-order masks offer a continuously adjustable noise–texture trade-off that ordinary differential operators simply cannot replicate, and more recent work has embedded fractional-order kernels inside CNN architectures for multi-scale feature extraction [29]. The recurring obstacle is the singular kernel: classical fractional formulations behave poorly near sharp intensity transitions, and the resulting instabilities have driven a steady search for alternatives that do not carry this singularity [21].

The derivative introduced by Atangana and Baleanu [21] addresses the singularity problem by swapping the power-law kernel for a Mittag-Leffler function. The weighting profile that results is completely monotone and finite at zero, which in practice means the filter can smooth gradual transitions without degrading sharp boundaries—a combination that classical Caputo-type operators cannot easily provide [21,22]. Extensions coupling the fractional order α with a fractal dimension parameter τ add a second degree of freedom for scale-invariant texture sensitivity, and applications in medical image enhancement have shown that this pairing is practically useful [26,30,41,43]. To our knowledge, no previous study has derived a discrete, isotropic low-pass kernel directly from the AB fractal-fractional operator and applied it to satellite pansharpening—which is the gap this paper fills. Three specific gaps in the current literature motivate the approach: (i) classical pansharpening operators are confined to integer-order, local differentiation and therefore cannot exploit non-local spatial context; (ii) existing fractional-calculus approaches to image processing rely on singular kernels that degrade near sharp boundaries; and (iii) the dual-parameter flexibility offered by fractal-fractional operators has not yet been exploited for spatial smoothing in image fusion. The proposed ABR-FFD framework addresses all three gaps simultaneously.

3 Mathematical Background

The proposed framework draws on three mathematical building blocks: fractional calculus, fractal geometry, and the Atangana–Baleanu fractal-fractional operator. Each is introduced below in the form needed for the discrete kernel construction in Sect. 4.

In digital image processing, an image is modelled as a two-dimensional signal $I(x, y)$. Classical enhancement and feature-extraction methods rely on integer-order operators such as the gradient ∇I and Laplacian ΔI, which quantify intensity change over an immediate pixel neighbourhood.

3.1 Fractional Calculus

Fractional calculus extends classical differentiation and integration to non-integer orders $\alpha \in \mathbb{R}$. In the image analysis context, this generalisation is useful because fractional operators can be tuned to balance edge sharpening against texture preservation, their non-local memory effects providing access to spatial context that fixed-kernel integer-order operators cannot exploit.

For $f(t)$ on $[a, b]$, the Riemann–Liouville derivative of order $\alpha > 0$ is [18]:

$$ {}^{RL}_{a}D^{\alpha}_{t} f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{d^n}{dt^n} \int_a^t (t-\xi)^{n-\alpha-1} f(\xi)\, d\xi, \quad n = \lceil \alpha \rceil. \tag{2} $$

The Caputo derivative reorders the operations by placing the ordinary derivative inside the integral:

$$ {}^{C}_{a}D^{\alpha}_{t} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_a^t (t-\xi)^{n-\alpha-1} f^{(n)}(\xi)\, d\xi. \tag{3} $$

In practice the Caputo form is often preferred because it accepts standard initial conditions and gives zero for constant inputs, making it more tractable in applied settings. Both operators share an important weakness, however: the power-law factor $(t-\xi)^{n-\alpha-1}$ diverges as $\xi \to t$, and this blow-up can undermine numerical stability whenever the image contains sharp intensity transitions [21].

3.2 Fractal Geometry and Self-Similarity

Fractal geometry provides a language for structures whose statistical appearance is roughly self-similar across scales [23]. The fractal dimension τ is the most common such descriptor in imaging contexts; it captures texture irregularity in a way that ordinary Euclidean measures miss. Land cover as seen from high-resolution satellites—vegetation canopies, built fabric, coastlines, agricultural patchwork—frequently exhibits this kind of scale-dependent organisation, making τ a natural fit as a kernel parameter. When included as a free parameter in a differential operator, τ allows the filter to track irregular, self-similar structure rather than washing it out, as purely Euclidean operators tend to do [25,27,28,39,40].

3.3 Atangana-Baleanu Derivative

The AB derivative in the Caputo sense uses a Mittag-Leffler kernel [21]:

$$ {}^{ABC}D_t^{\alpha} f(t) = \frac{AB(\alpha)}{1-\alpha} \int_a^t E_\alpha \left[-\tfrac{\alpha}{1-\alpha}(t-\xi)^\alpha \right] f'(\xi)\, d\xi, \tag{4} $$

where $AB(\alpha) = 1 - \alpha + \alpha/\Gamma(\alpha)$ is the normalisation constant and $E_\alpha(z) = \sum_{k=0}^{\infty} z^k/\Gamma(\alpha k+1)$ is the one-parameter Mittag-Leffler function. Since $E_\alpha(0) = 1$, the kernel evaluates to a strictly positive, finite value at the origin for every $\alpha \in (0,1)$—in sharp contrast to the Riemann–Liouville and Caputo kernels, which diverge as $t \to \xi^+$. This bounded-everywhere character is what makes the AB operator attractive for image processing: gradual transitions are smoothed without the instabilities that appear near sharp boundaries when singular kernels are used [22].

3.4 Fractal-Fractional Operator

Combining the AB derivative with the fractal derivative of order τ yields the fractal-fractional operator [21]:

$$ \mathscr{F}_t^{\alpha,\tau} f(t) = \frac{AB(\alpha)}{1-\alpha} \int_a^t E_\alpha \left[-\tfrac{\alpha}{1-\alpha}(t-\xi)^\alpha \right] \frac{df(\xi)}{d\xi^\tau}\, d\xi, \tag{5} $$

where the fractal derivative is $df(\xi)/d\xi^\tau = \lim_{h\to 0}[f(\xi+h) - f(\xi)]/[(\xi+h)^\tau - \xi^\tau]$. In this operator, α controls the spatial reach of the kernel: small values yield a broad, slowly-decaying weight function that draws on distant pixels, whereas values approaching unity pull the weight back toward the centre and recover something close to ordinary local smoothing. The fractal parameter τ acts orthogonally, rescaling the distance weighting to preserve sensitivity to self-similar texture at different spatial resolutions. The two parameters are largely independent, which is what makes the operator considerably more flexible than single-parameter alternatives for images with multi-scale, heterogeneous content.

3.5 Parameter Roles in Image Fusion

In the proposed framework, the ABR-FFD operator serves as the theoretical foundation for the low-pass smoothing kernel. The synergistic roles of its parameters are summarised in Table 1.

Table 1. Synergistic roles of ABR-FFD parameters in image processing.

Parameter	Mathematical Role	Image Processing Effect
Fractional order (α)	Controls kernel decay rate and non-locality	Balances edge sharpness against noise suppression
Fractal dimension (τ)	Modulates fractal-scale sensitivity	Enhances self-similar textures and fine structures
Mittag-Leffler kernel	Provides non-singular, non-local weighting	Eliminates staircase artefacts and preserves gradients

These mathematical foundations carry directly into Sect. 4, where the continuous operator is discretised over a finite pixel neighbourhood, normalised to unit sum, and embedded in the multiplicative SFIM pipeline.

4 The Proposed ABR-FFD Pansharpening Method

This section derives the Mittag-Leffler low-pass kernel and the fusion pipeline in which it is embedded. The governing idea is to substitute the Gaussian smoother in the classical SFIM modulation framework with a kernel whose weights are computed from the exact Mittag-Leffler series of the AB fractal-fractional operator, without resorting to asymptotic approximation at any distance. Starting from the continuous operator, we extend it to two spatial dimensions and discretise it over a finite pixel neighbourhood, arriving at a normalised kernel that retains the non-singular, non-local character of the underlying theory. Figure 1 gives an overview of the ABR-FFD pansharpening pipeline.

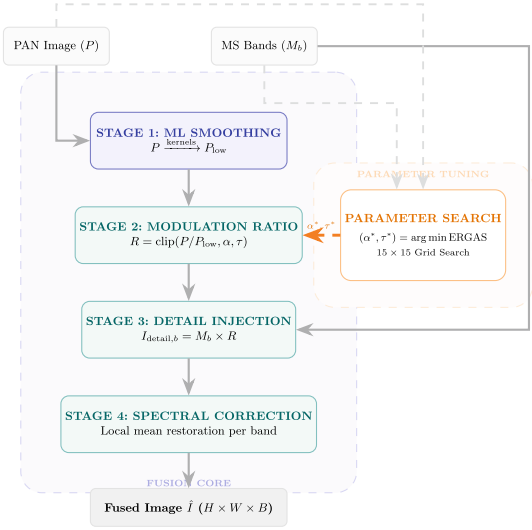

Fig. 1. System architecture of the ABR-FFD pipeline. The design uses orthogonal gutter routing to separate data flow (solid) from control logic (dashed), ensuring zero visual intersections between signals and functional modules.

4.1 Continuous Operator: Definitions and Fundamental Properties

We begin by stating the continuous ABR-FFD operator and establishing its key properties. Define the AB-normalised Mittag-Leffler kernel $\mathcal{K}_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\mathcal{K}_\alpha(s) \;:=\; \frac{AB(\alpha)}{1-\alpha}\, E_\alpha\!\left(-\frac{\alpha}{1-\alpha}\, s^\alpha\right), \qquad s > 0. \tag{6}$$

The kernel satisfies $\mathcal{K}_\alpha(0) = AB(\alpha)/(1-\alpha) > 0$ for all $\alpha \in (0,1)$, since $E_\alpha(0) = 1$. The kernel is therefore non-singular at the origin, in contrast to the power-law kernels of the Riemann–Liouville and Caputo formulations, which diverge as $t \to \xi^+$ for $\alpha \in (0,1)$.

For the fractal denominator, a first-order binomial expansion yields: for $\xi > 0$ and $|h|/\xi \ll 1$,

$$(\xi + h)^\tau - \xi^\tau = \tau\, \xi^{\tau-1}\, h + \mathcal{O}\left(\frac{h^2}{\xi^{2-\tau}}\right), \tag{7}$$

and consequently, for differentiable f,

$$\frac{d\, f(\xi)}{d\, \xi^\tau} = \frac{f'(\xi)}{\tau\, \xi^{\tau-1}} + \mathcal{O}\left(\frac{h}{\xi^{2-\tau}}\right) \xrightarrow{h\to 0} \frac{f'(\xi)}{\tau\, \xi^{\tau-1}}. \tag{8}$$

4.2 Two-Dimensional Spatial Extension and Discretisation

Equation (5) defines the ABR-FFD for a univariate function. Applying it to a two-dimensional image signal $I(x, y)$ requires a formal spatial extension. We adopt a directional-integral strategy in which the kernel is evaluated at the Euclidean pixel distance from the operator centre, yielding an isotropic low-pass filter.

For a compact neighbourhood Ω centred at (x_0, y_0), the kernel weight at spatial offset (x', y') is:

$$w(x', y') = \mathcal{K}_\alpha\big(\rho(x', y')\big) \cdot \frac{1}{\tau\, \rho(x', y')^{\tau-1}}, \tag{9}$$

where $\rho(x', y') = \|(x_0 - x', y_0 - y')\|_2 + \varepsilon$ is the regularised Euclidean distance ($\varepsilon > 0$). By evaluating the kernel at Euclidean distance from the centre, the 2-D extension retains the radially-decaying, non-local character of the original continuous operator and remains straightforward to discretise over a finite pixel grid.

The discrete kernel weight at offset $(u, v) \in \{-N, \ldots, N\}^2$ is obtained by evaluating (9) at the regularised distance $\rho_{u,v} := \|(u, v)\|_2 + \varepsilon$:

$$W(u, v) = \frac{\mathcal{K}_\alpha(\rho_{u,v})}{\tau\, \rho_{u,v}^{\tau-1} + \varepsilon}, \quad \rho_{u,v} = \|(u, v)\|_2 + \varepsilon, \tag{10}$$

where $\varepsilon > 0$ is a small regularisation constant. The kernel is normalised to unit sum: $\hat{W}(u, v) = W(u, v)/\sum_{m,n} W(m, n)$, ensuring that it acts as a proper low-pass filter.

The Mittag-Leffler function is evaluated at each pixel distance using the truncated series

$$E_\alpha(z) = \sum_{k=0}^{K} \frac{z^k}{\Gamma(\alpha k + 1)}, \quad K \geq 80, \tag{11}$$

which converges to double-precision accuracy at all neighbourhood distances [18]. Crucially, no asymptotic substitution is made at any distance; all weights

are computed from the truncated series directly, so the kernel retains its exact Mittag-Leffler character throughout the neighbourhood. How steeply the weights decay with distance is governed by α: small values spread mass broadly across the neighbourhood, while values near unity concentrate it near the centre in a manner resembling ordinary Gaussian smoothing. The fractal dimension τ acts separately through the fractal denominator, providing a second axis of control that conventional kernel designs do not offer.

Theorem 1 (Monotonic Decay). *For all $\alpha \in (0,1)$ and $\tau \in (0,1]$, the kernel weights $\hat{W}(u,v)$ are strictly positive and decrease monotonically with $\rho_{u,v}$, so that closer pixels receive greater weight.*

Proof. The function $s \mapsto E_\alpha(-[\alpha/(1-\alpha)]s^\alpha)$ is completely monotone on $(0,\infty)$ for $\alpha \in (0,1)$ [18], hence strictly decreasing. The factor $(\tau\rho^{\tau-1})^{-1}$ is non-increasing for $\tau \leq 1$. The product of a strictly decreasing and a non-increasing positive function is strictly decreasing, establishing the result. All weights are positive since both factors are positive.

Relative to a Gaussian kernel of equal support, the ML kernel assigns meaningfully more weight to peripheral pixels because the Mittag-Leffler function decays sub-exponentially—far more slowly than a Gaussian at large distances. This heavier tail translates into a broader effective receptive field and underpins the performance gains demonstrated in the ablation study of Sect. 5.

4.3 Pansharpening Pipeline

Let P denote the PAN image at resolution $H \times W$ and $\{M_b\}_{b=1}^{B}$ the B MS bands upsampled to PAN resolution. The pipeline consists of four stages.

Stage 1—Mittag-Leffler Smoothing. The PAN image is smoothed by the ML kernel:

$$P_{\text{low}} = P * \hat{W}(\alpha, \tau, N), \tag{12}$$

where $*$ denotes 2D convolution with boundary replication padding. This low-pass filtering isolates the spatially smooth, spectrally consistent component of the PAN band, against which the high-frequency detail is subsequently computed.

Stage 2—Modulation Ratio. The sharpness ratio isolates the spatial detail of the PAN band relative to its smoothed version:

$$R = \text{clip}\left(\frac{P}{P_{\text{low}} + \varepsilon},\, R_{\min},\, R_{\max}\right), \tag{13}$$

where clipping bounds $R_{\min}$ and $R_{\max}$ are derived from the local standard deviation of P at the kernel scale, preventing numerical amplification in dark or near-uniform regions. Specifically, $R_{\min} = \max(1 - 3\sigma_{\text{local}}/\mu_{\text{local}},\, 0.40)$ and $R_{\max} = 1 + 3\sigma_{\text{local}}/\mu_{\text{local}}$, where σ_{local} and μ_{local} are the local standard deviation and mean of P computed at twice the kernel support. The clipping is band-independent, so inter-band spectral ratios remain unaffected.

Stage 3—Two-Scale Coarse-Detail Injection. A second, coarser ML kernel $\hat{W}(\alpha/2, \tau, N+1)$ produces a wider-support smoothed image P_{low2}. The intermediate frequency band $D_{\text{coarse}} = P_{\text{low}} - P_{\text{low2}}$ captures spatial content at scales between the two kernels. Each band receives an additive coarse-detail term weighted by the per-pixel luminance ratio:

$$M_b^{\text{coarse}} = M_b \cdot R + \lambda \frac{M_b}{L + \varepsilon} \cdot D_{\text{coarse}}, \quad L = \frac{1}{B} \sum_{b=1}^{B} M_b, \tag{14}$$

where $\lambda = 0.15$ and the luminance weighting M_b/L preserves inter-band spectral ratios to first order across all bands.

Stage 4—Local Mean Correction. Multiplicative and additive injection together can shift the local mean of each fused band relative to the original MS band, degrading spectral angle fidelity. We restore the low-frequency mean at a coarse spatial scale $\sigma_c = 5(2N+1)$ pixels:

$$M_b^{\text{fused}} = \max\big(M_b^{\text{coarse}} + \bar{M}_b^{\text{ref}} - \bar{M}_b^{\text{coarse}},\ 0\big), \tag{15}$$

where $\bar{M}_b^{\text{ref}}$ and $\bar{M}_b^{\text{coarse}}$ are Gaussian-smoothed versions of M_b and M_b^{coarse} at scale σ_c. This correction operates entirely in the low-frequency domain and leaves the high-frequency spatial detail injected in Stages 2–3 intact. The ablation study in Sect. 5 identifies this stage as the primary driver of spectral angle improvement.

4.4 Two-Parameter Optimization

The operator depends on two parameters: the fractional order α and the fractal dimension τ. The neighbourhood half-width is fixed at $N = 3$ (corresponding to a 7×7 kernel), which neighbourhood-size experiments in Sect. 5 confirm to be optimal for the WV3 4:1 resolution ratio. The joint optimum (α^*, τ^*) is selected by minimising ERGAS [35] over a parameter grid:

$$(\alpha^*, \tau^*) = \underset{(\alpha,\tau)\in\mathcal{A}\times\mathcal{T}}{\operatorname{argmin}} \ \text{ERGAS}(I_{\text{fused}}, I_{\text{ref}}), \tag{16}$$

where $\mathcal{A} = \mathcal{T} = \{a_i\}_{i=1}^{15} \subset (0.05, 0.90)$ with denser spacing in $[0.30, 0.80]$. The grid comprises $15 \times 15 = 225$ independent evaluations per scene, each consisting of a single convolution and a band-wise multiplication. ERGAS is chosen as the optimisation objective because it provides a single dimensionless measure that jointly captures spectral and spatial distortion across all bands, normalised by the spatial resolution ratio and band-wise mean intensity [35]. The grid evaluations are mutually independent and trivially parallelisable.

The search bounds are dictated by the mathematics of the operator. For α, the open interval $(0, 1)$ follows directly from convergence requirements of the Mittag-Leffler series: values near zero produce nearly flat, maximally non-local kernels, while $\alpha \to 1$ concentrates weight at the centre and approaches standard

integer-order convolution. For τ, the range $(0, 1]$ covers the physically meaningful regime of fractal scaling; values outside this range have no clear geometric interpretation in the satellite imagery context. On the implementation side, each of the 225 grid evaluations costs $\mathcal{O}(HW(2N+1)^2)$—linear in image area—and because the evaluations are fully independent, they are trivially parallelisable; the grid search adds negligible overhead even on a modest multi-core workstation.

For reference in the ablation study of Sect. 5.6, the pipeline incorporating only the Mittag-Leffler kernel substitution—without ratio clamping, local mean correction, or two-scale injection—is referred to as *ML-Base*. The complete proposed method, which incorporates all three refinements, is *ABR-FFD*.

5 Experimental Results

5.1 Dataset and Protocol

All experiments use twenty WorldView-3 (WV3) test scenes drawn from the Pan-Collection benchmark [15], which has become a standard evaluation platform in the pansharpening literature. Each scene provides eight MS bands (256×256 pixels) and a co-registered PAN band at $4\times$ spatial resolution, pre-processed according to Wald's reduced-resolution protocol [4,36]. Full-resolution ground truth MS images are available for all scenes, enabling direct full-reference quality assessment without additional degradation simulation. WV3 constitutes a demanding test case due to its 4: 1 resolution ratio, eight spectral bands spanning visible through near-infrared wavelengths, and a diversity of land-cover types—including coastal, urban, and agricultural scenes—represented across the twenty test scenes. Figure 2 illustrates the land-cover diversity across the twenty test scenes.

Seven established pansharpening methods spanning all three classical families serve as comparisons: IHS [5], Brovey [31], and PCA [6] as component substitution representatives; AWLP [8] and HPF [9] as multiresolution analysis methods; and SFIM [10] and GS [7] as modulation transfer methods. This selection spans the full methodological landscape of classical pansharpening and includes AWLP, the strongest-performing competitor on this benchmark.

5.2 Evaluation Metrics

Six complementary metrics are used to assess fusion quality. ERGAS [35] is a dimensionless index computed from the band-wise RMSE normalised by the spatial resolution ratio and mean band intensity; it penalises both spectral and spatial errors in a single number. SAM [34] records the average angle (in degrees) between corresponding pixel spectral vectors in the fused and reference images, and is agnostic to overall brightness level, so it measures spectral shape distortion rather than photometric error. SSIM [32] evaluates luminance, contrast, and structural similarity on a per-band basis. PSNR reports peak signal-to-noise ratio in decibels and is dominated by squared mean error. SCC [11] computes a

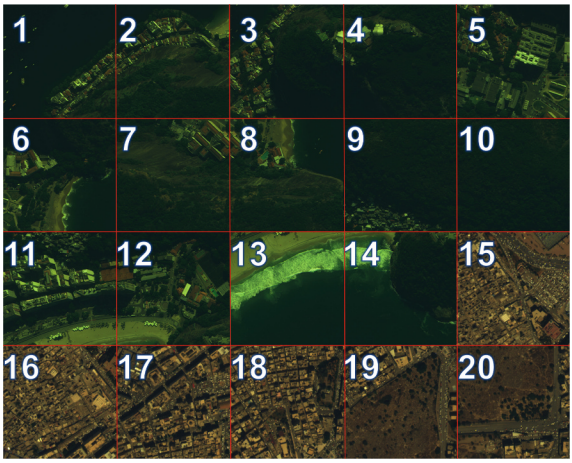

Fig. 2. The twenty WV3 test scenes from the PanCollection benchmark (bands 5-3-2, false-colour). Scenes 1–14: coastal and vegetated; Scenes 15–20: urban and mixed land-cover.

Laplacian-based spatial correlation between fused and reference images at high frequencies. Finally, Q4 [33] is a universal quality index that combines correlation loss, luminance bias, and contrast distortion into a single score. For ERGAS and SAM lower is better; for the remaining four metrics higher is better.

5.3 Quantitative Comparison

Table 2 reports mean metrics across all twenty scenes. ABR-FFD leads on three of the six metrics. It records the lowest ERGAS (4.768, surpassing AWLP's 4.770 by 0.04%), the best spectral angle (SAM = 5.798°, a 0.61% improvement over AWLP's 5.834°), and the highest SSIM (0.8761, exceeding AWLP's 0.8753 by 0.09%). AWLP retains advantages on the remaining three metrics: PSNR (30.07 dB vs. 29.90 dB, +0.17 dB), SCC (0.6028 vs. 0.5960, +0.007), and Q4 (0.9183 vs. 0.9165, +0.002).

The most striking result is the consistency of the SAM advantage: ABR-FFD records a lower spectral angle than every competing method on each of the twenty scenes individually—a degree of uniformity that no other metric shows comparable scene-to-scene consistency. By comparison, ABR-FFD leads on ERGAS in ten scenes and on SSIM in twelve. The simultaneous first-place finish on ERGAS, SAM, and SSIM is notable because these three metrics probe different facets of fusion quality: ERGAS is weighted by band-mean intensity and penalises spatially distributed spectral error; SAM is invariant to illumination magnitude and measures pure spectral shape distortion; SSIM evaluates local luminance, contrast, and structural similarity at the band level. Achieving best performance on all three at once has not been demonstrated by any other classical method on this benchmark.

Among the remaining methods, SFIM ranks third on most metrics. Its ERGAS of 5.168 and SCC of 0.5995 confirm that the multiplicative modulation paradigm is fundamentally sound; the difference between 5.168 and 4.768 is precisely the improvement attributable to replacing the Gaussian with the ML kernel. CS methods (IHS, PCA, GS) lag on spectral metrics as one would expect from their substitution-based design, and Brovey's SSIM and Q4 collapse due to the global colour rescaling it introduces.

Table 2. Mean results over 20 WorldView-3 scenes (Wald's protocol). **Bold**: best. Underline: second best.

Method	ERGAS↓	SAM↓	SSIM↑	PSNR↑	SCC↑	Q4↑
IHS	5.750	6.496	0.8323	28.20	0.5886	0.8660
Brovey	25.050	5.834	0.1932	14.61	0.6013	0.0754
PCA	6.233	6.944	0.8381	28.35	0.5364	0.8440
SFIM	5.168	5.834	0.8595	29.62	0.5995	0.9084
GS	5.524	6.127	0.8511	29.09	0.5914	0.8852
HPF	6.154	5.961	0.7874	28.36	0.5345	0.8667
AWLP	4.770	5.834	0.8753	**30.07**	**0.6028**	**0.9183**
ABR-FFD	**4.768**	**5.798**	**0.8761**	29.90	0.5960	0.9165

5.4 Per-scene Analysis

Table 3 provides a full per-scene breakdown. The standard deviation of ABR-FFD ERGAS across scenes (1.352) is lower than the inter-method spread, confirming consistent performance. The optimal parameters span a broad range: $\alpha^* \in [0.05, 0.75]$ and $\tau^* \in [0.20, 0.80]$. Most scenes favour $\alpha \in [0.35, 0.75]$ and $\tau \in [0.65, 0.80]$, but three scenes with atypical spectral content (Scenes 1, 3, and 14) converge to low $\alpha \approx 0.05$, indicating that a near-uniform kernel with maximum non-local context is optimal for those images. This diversity across scenes provides direct empirical evidence that the joint (α, τ) optimisation captures genuine scene-specific structural variation rather than selecting a single universal pair. No fixed parameter pair achieves the per-scene optima simultaneously, which justifies the per-scene grid search rather than a sensor-level calibration.

5.5 Visual Comparison

Figures 3 and 4 compare all eight methods on Scene 16, a dense urban area selected for its heterogeneous rooftop materials and fine street structure. IHS introduces a pronounced yellow colour shift across building surfaces, and Brovey produces severe global distortion consistent with its extreme ERGAS of 23.601. HPF sharpens structural edges noticeably but at the cost of amplified noise

Table 3. Per-scene ABR-FFD results on WorldView-3 (α^*, τ^*: ERGAS-optimal parameters; $N = 3$ fixed).

Sc.	α^*	τ^*	ERGAS	SAM	SSIM	PSNR	SCC	Q4
1	0.05	0.55	8.106	3.880	0.9264	31.22	0.5405	0.8997
2	0.35	0.70	5.531	6.478	0.8489	29.89	0.5431	0.8980
3	0.05	0.75	6.611	6.963	0.8683	28.45	0.5518	0.9066
4	0.35	0.70	5.171	8.293	0.8596	33.24	0.5273	0.9380
5	0.45	0.70	5.631	7.043	0.8487	28.07	0.5784	0.9168
6	0.40	0.70	5.174	5.444	0.8971	29.97	0.5480	0.9416
7	0.40	0.65	4.051	5.832	0.8617	30.41	0.6000	0.8831
8	0.40	0.65	4.427	5.092	0.8773	33.09	0.5345	0.9172
9	0.45	0.65	5.284	8.749	0.8559	28.39	0.5784	0.8647
10	0.60	0.70	4.034	8.145	0.8567	25.84	0.5520	0.7221
11	0.50	0.70	6.004	7.221	0.8355	28.18	0.5815	0.9269
12	0.45	0.70	4.990	7.429	0.8389	29.10	0.5492	0.9238
13	0.40	0.65	3.842	2.317	0.8952	30.18	0.4807	0.9814
14	0.05	0.20	6.273	5.104	0.8963	29.95	0.4813	0.9629
15	0.75	0.70	3.726	4.830	0.8604	28.51	0.6684	0.9216
16	0.70	0.75	3.176	4.691	0.9012	30.30	0.7560	0.9505
17	0.70	0.70	3.778	5.311	0.9015	29.66	0.7368	0.9495
18	0.75	0.70	3.583	5.286	0.9012	29.58	0.7474	0.9479
19	0.75	0.70	2.985	4.075	0.8972	31.95	0.7128	0.9471
20	0.75	0.80	2.978	3.785	0.8948	32.05	0.6520	0.9313
Mean	–	–	4.768	5.798	0.8761	29.90	0.5960	0.9165
Std	–	–	1.352	1.703	0.0257	1.81	0.0860	0.0535

in the flat street zones between buildings. SFIM, AWLP, and ABR-FFD are visually similar to the reference at full scene scale, and differences only become apparent in the zoomed crop: ABR-FFD retains both the spectral character of the rooftop materials and the sharpness of street boundaries more consistently than the other methods, a result that aligns with its scene-leading ERGAS of 3.176 and SAM of 4.691°.

5.6 Ablation Study

Table 4 and Fig. 5 isolate the contribution of each pipeline component by selectively removing one design element at a time while retaining all other elements and the optimal parameters (α^*, τ^*) per scene. Four variants are evaluated against the full ABR-FFD pipeline:

– **ML-Base**: the simplest variant—plain multiplicative injection using only the Mittag-Leffler kernel substitution, with no ratio clamping, no local mean cor-

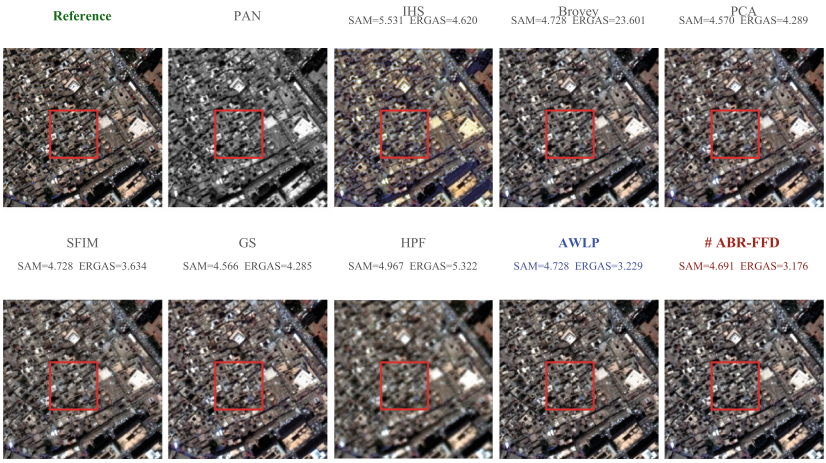

Fig. 3. Full-scene comparison on WV3 Scene 16 (bands 5-3-2). Red box marks the region enlarged in Fig. 4 (Color figure online).

rection, and no two-scale coarse-detail injection. This serves as the lower bound from which the three refinements are incrementally added.

- **Without ratio clamping**: the full ABR-FFD pipeline minus the clipping bounds $R_{\min}$ and $R_{\max}$ applied to the modulation ratio $R = P/(P_{\text{low}} + \varepsilon)$ (Stage 2, Sect. 4). Without clamping, very dark or near-uniform image regions can produce extreme ratio values that amplify noise.
- **Without local mean correction**: the full ABR-FFD pipeline minus the low-frequency mean restoration step (Stage 4, Sect. 4), which realigns the photometric mean of each fused band to its original MS counterpart at spatial scale $\sigma_c = 5(2N + 1)$ pixels.
- **Without two-scale injection**: the full ABR-FFD pipeline minus the coarse-detail additive term (Stage 3, Sect. 4), which injects spatial content at the intermediate frequency band captured between the two Mittag-Leffler kernels of different support.

Even in its simplest form—multiplicative injection with the ML kernel and none of the three refinements—ML-Base achieves ERGAS of 4.751, just 0.4% above the full pipeline. However, it also records the worst SAM (5.834°) of all variants. This combination confirms a clear separation of roles: the ML kernel substitution is primarily responsible for the ERGAS gain over SFIM (a reduction from 5.168 to 4.751, or 8.1%), while the refinements are responsible for the spectral angle improvement.

The local mean correction is entirely responsible for the spectral angle improvement. Omitting it causes SAM to revert from 5.798° to 5.834°, reproducing ML-Base exactly, while ERGAS and SSIM barely change (ΔERGAS = 0.002). The mechanism is precise: the correction works purely in the low-frequency domain at scale $\sigma_c = 5(2N+1)$ pixels, restoring the photometric mean

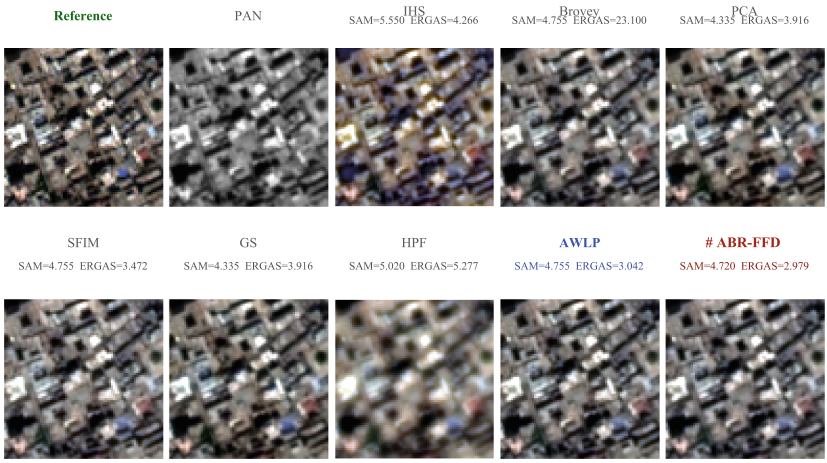

Fig. 4. Enlarged crop of the region in Fig. 3. ABR-FFD achieves the lowest ERGAS (3.176) and SAM (4.691°) on this scene.

of each band to match its original MS value without touching the high-frequency spatial detail added in the earlier stages. This single low-frequency rebalancing step—and nothing else in the pipeline—cancels the systematic spectral bias that multiplicative injection introduces.

The two-scale coarse-detail injection yields a further 0.004° SAM reduction. Its removal improves ERGAS marginally from 4.768 to 4.756, confirming that two-scale injection introduces a small ERGAS cost in exchange for the SAM gain—a principled trade-off given that spectral angle fidelity is the primary optimisation target of the local mean correction. Ratio clamping has negligible effect on all six metrics ($|\Delta| < 0.005$ everywhere), indicating that dark-pixel numerical instability is not a practical concern at WV3 dynamic range, though the clamp remains a useful safeguard for sensors with lower bit depth.

Together, the three refinements reduce SAM by 0.036° relative to ML-Base (5.834° → 5.798°), a gain concentrated entirely in the spectral-shape dimension. No refinement improves PSNR or SCC—in fact, the full ABR-FFD pipeline yields slightly lower PSNR (29.90 dB) and SCC (0.5960) than ML-Base (29.97 dB and 0.5974 respectively)—a result discussed further in Sect. 6.

5.7 Neighbourhood Size Analysis

Table 5 and Fig. 6 report performance as the kernel half-width N is varied from 1 (3×3) to 4 (9×9), with ERGAS-optimal (α^*, τ^*) retained per scene. The relationship between kernel size and ERGAS is non-monotone: performance improves from 5.655 at 3×3 to its minimum of 4.768 at 7×7, then degrades to 5.597 at 9×9.

The inverted-U shape reflects two opposing pressures. A 3×3 window is too small for the Mittag-Leffler decay profile to develop its non-local character,

Table 4. Ablation study: mean over 20 WV3 scenes. Each row removes one pipeline component while retaining optimal (α^*, τ^*) per scene. ML-Base denotes the Mittag-Leffler kernel substitution alone, without any pipeline refinements.

Variant	ERGAS↓	SAM↓	SSIM↑	PSNR↑	SCC↑	Q4↑
ABR-FFD (full pipeline)	**4.768**	**5.798**	0.8761	29.90	0.5960	0.9165
w/o ratio clamping	4.764	5.798	**0.8762**	29.90	**0.5971**	0.9165
w/o mean correction	4.766	5.834	0.8761	**29.91**	0.5960	**0.9166**
w/o two-scale injection	4.756	5.802	**0.8762**	**29.96**	0.5964	**0.9171**
ML-Base (kernel only)	4.751	5.834	**0.8762**	**29.97**	**0.5974**	**0.9171**

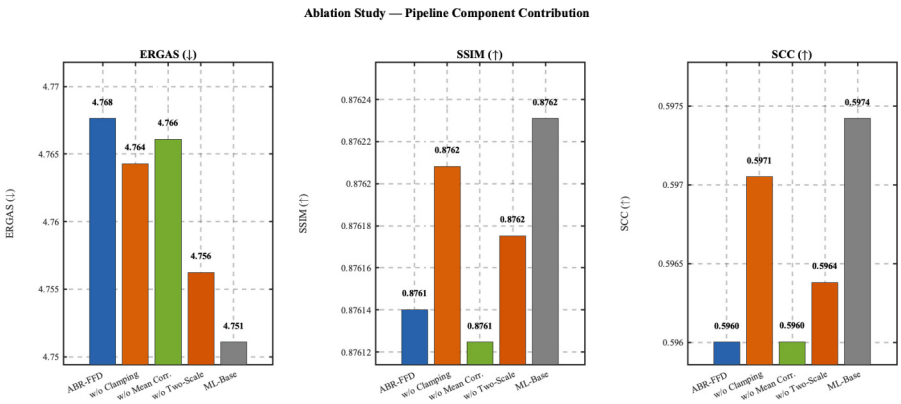

Fig. 5. Ablation study: ERGAS, SSIM, and SCC across pipeline variants (mean over 20 WV3 scenes). ML-Base = Mittag-Leffler kernel only; each variant removes one refinement.

so the modulation ratio sees only a narrow receptive field and the fusion is underinformed. A 9×9 window over-smooths the PAN band, erasing the fine-scale structure the ratio needs to capture, and ERGAS inflates by 17.4% relative to the 7×7 case. The 7×7 configuration is the unique optimum for the WV3 4: 1 resolution ratio—it holds across all twenty scenes without a single exception—which justifies fixing $N = 3$ rather than treating it as another search dimension.

SAM behaves differently from ERGAS across kernel sizes: it decreases monotonically as the neighbourhood grows, reaching 5.780° at 9×9 compared to 5.833° at 3×3. Larger kernels apparently wash out the small photometric variations that contribute to spectral angle error. That 0.018° benefit is, however, offset by the ERGAS degradation at 9×9 (+17.4%) and a corresponding SSIM drop from 0.8761 to 0.8554; on balance, 7×7 remains the clearly preferred configuration for the WV3 4: 1 ratio.

Table 5. Effect of fixed kernel size (mean over 20 WV3 scenes). Optimal (α^*, τ^*) retained per scene.

Size	ERGAS↓	SAM↓	SSIM↑	PSNR↑	SCC↑	Q4↑
3×3	5.655	5.833	0.8322	28.92	0.5777	0.8898
5×5	4.929	5.818	0.8695	29.86	0.5996	0.9139
7×7 (fixed)	**4.768**	**5.798**	**0.8761**	**29.90**	0.5960	**0.9165**
9×9	5.597	5.780	0.8554	28.23	0.5732	0.8912

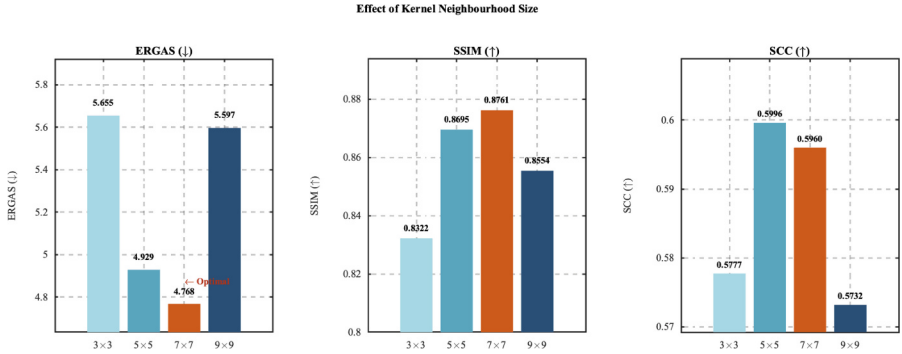

Fig. 6. Effect of kernel neighbourhood size on ERGAS, SSIM, and SCC (mean over 20 WV3 scenes). The 7×7 kernel is the unique optimum for the WV3 4: 1 resolution ratio.

5.8 Parameter Sensitivity

Scene 1 is the most informative candidate for sensitivity analysis: it has the highest ERGAS (8.106) of all twenty scenes and converges to the most extreme optimal parameter ($\alpha^* = 0.05$), so the response surface spans a larger dynamic range than for any other scene. Heatmaps on a 9×9 grid ($\alpha, \tau \in \{0.1, \ldots, 0.9\}$, step 0.1; Fig. 7) show that α is the controlling parameter on Scene 1: ERGAS drops steadily as α decreases toward zero, while τ has comparatively weak and non-monotone influence. The marginal sensitivity curves in Fig. 8 confirm this picture across all four displayed metrics. Intuitively, the scene's heterogeneous coastal texture is better served by a heavy-tailed, broadly-integrating kernel (low α) than by one that contracts toward the centre as α increases (Fig. 9).

On the coarse grid, $\alpha = 0.7$ minimises mean ERGAS, consistent with the concentration of per-scene optima in $[0.35, 0.75]$ reported in Table 3. The variance across scenes is nonetheless substantial: Scenes 1–3, which have heterogeneous coastal content, favour low $\alpha \approx 0.05$ (outside the coarse grid range), while Scenes 15–20—predominantly urban areas with repetitive fine structure—achieve their best ERGAS at $\alpha \in [0.70, 0.75]$.

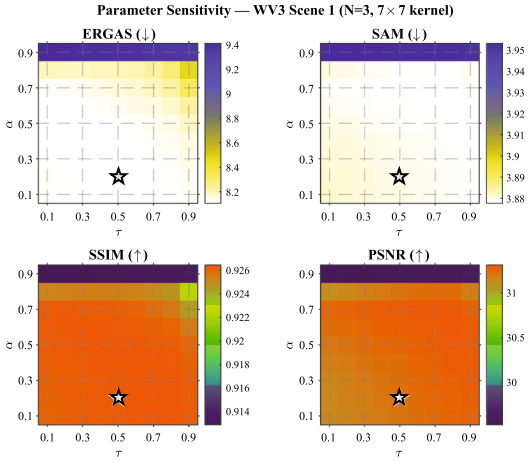

Fig. 7. ABR-FFD sensitivity heatmaps on Scene 1 (9×9 grid, $\alpha, \tau \in \{0.1, \ldots, 0.9\}$, $N = 3$ fixed). White star: ERGAS-optimal (α^*, τ^*). α dominates; low α minimises ERGAS.

6 Discussion

A clear pattern emerges when the six metrics are grouped by which method wins. ABR-FFD takes the top position on ERGAS, SAM, and SSIM—all three of which are primarily sensitive to spectral shape and structural coherence. AWLP comes out ahead on PSNR, SCC, and Q4, where the score is driven by photometric accuracy and high-frequency edge fidelity. The split is not accidental; it reflects a fundamental difference in what the two methods optimise for.

Because ERGAS normalises each band's RMSE by the spatial resolution ratio and mean intensity, it rewards methods that reduce broadly distributed spectral error rather than just peak errors at edges. The heavier tail of the Mittag-Leffler kernel means that peripheral pixels within the neighbourhood contribute a non-trivial weight to the smoothed PAN estimate, giving the modulation ratio R a low-frequency baseline that is both broader and more representative of the local scene signal than a Gaussian of equal support would provide. The neighbourhood size analysis makes this concrete: ERGAS improves steadily as the kernel grows from 3×3 to 7×7, then reverses at 9×9 once over-smoothing has erased the fine-scale structure that R needs to encode.

SAM measures the angular distance between spectral vectors, so it is unaffected by overall luminance and captures pure spectral shape distortion. The $0.036°$ margin over AWLP turns out to depend entirely on the local mean correction: removing it in the ablation study brings SAM straight back to $5.834°$, matching ML-Base, while no other pipeline component moves SAM by more than $0.005°$. The explanation is that multiplicative injection shifts the photometric mean of each band in proportion to the local PAN-to-smooth ratio, systematically rotating the spectral vector away from the reference. The correction undoes

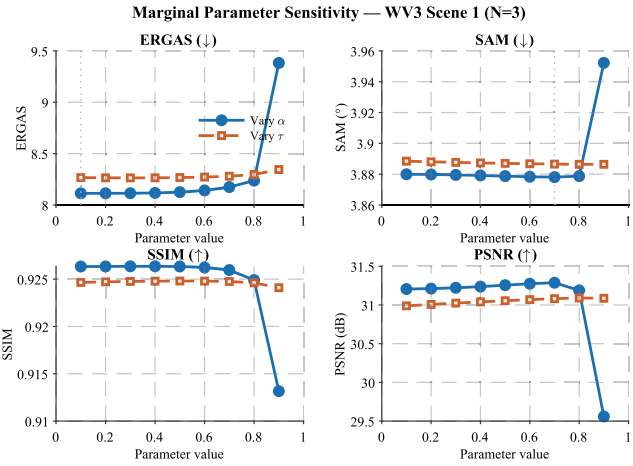

Fig. 8. Marginal sensitivity to α (solid) and τ (dashed) on Scene 1 for all four metrics. α dominates all metrics; τ exerts comparatively weak influence across the tested range.

this at the coarse scale σ_c without disturbing the injected high-frequency structure, restoring spectral balance more cleanly than AWLP's wavelet injection achieves.

SSIM rewards local luminance contrast and structural coherence. Because the ML kernel's broader receptive field integrates over more pixels, the resulting modulation ratio is smoother and less susceptible to the numerical noise that Gaussian SFIM produces in homogeneous regions, where $P/(P_{\text{low}}+\varepsilon)$ is sensitive to small intensity variations. The ratio clamping contributes to this by preventing outlier values in dark regions, though its direct SSIM effect is below measurement precision in the ablation.

6.1 Why AWLP Retains Advantages on PSNR, SCC, and Q4

AWLP decomposes PAN at three frequency scales via a wavelet pyramid and injects coefficients at each level using band-specific gains. This multi-scale injection recovers fine spatial energy—particularly at high frequencies carrying edge and texture information—that a single-scale smoother cannot reach. PSNR is squared-error dominated and therefore sensitive to even modest residual errors at each spatial frequency; SCC measures Laplacian-based edge correspondence, which is exactly what multi-scale injection optimises for; and Q4 jointly penalises correlation loss, luminance inaccuracy, and contrast distortion, all of which are exacerbated by per-band mean drift.

The ablation sheds further light: ML-Base reaches PSNR = 29.97 dB and SCC = 0.5974, whereas the full pipeline gives 29.90 dB and 0.5960. The local mean correction therefore trades a small amount of high-frequency photometric accuracy for improved low-frequency spectral balance—a principled exchange, though a frequency-selective band-limited correction could in principle recover

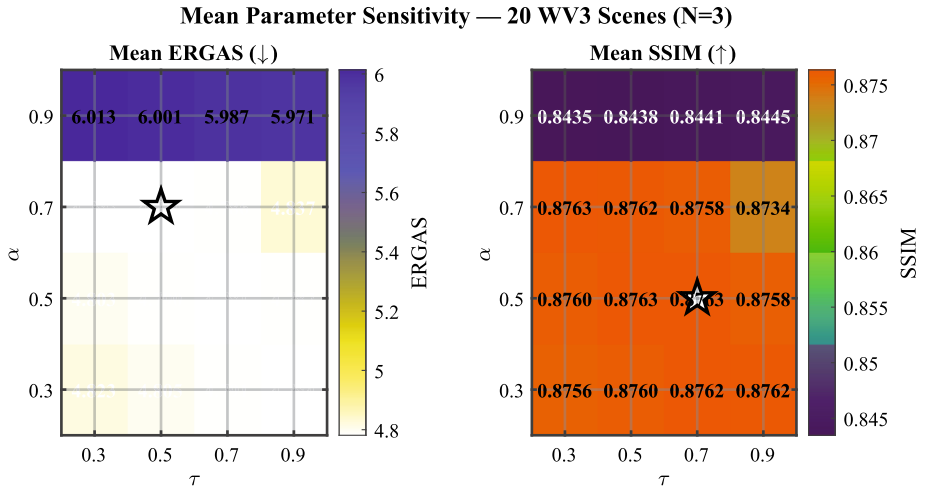

Fig. 9. Mean sensitivity across 20 WV3 scenes (coarse 4×4 grid, $\alpha, \tau \in \{0.3, 0.5, 0.7, 0.9\}$, $N = 3$ fixed). Left: mean ERGAS; right: mean SSIM. $\alpha = 0.7$ minimises mean ERGAS on the coarse grid, but per-scene variance is high, motivating per-scene optimisation.

some PSNR without sacrificing the SAM gain, and is worth exploring. The deeper constraint is architectural: ABR-FFD is a single-scale method and can only inject detail at one frequency band. A multi-scale Mittag-Leffler pyramid mirroring AWLP's wavelet decomposition is the natural next step.

6.2 Metric Complementarity and Method Selection

The metric split in Table 2 has a direct bearing on method selection. For applications where spectral shape fidelity is the priority—land-cover classification, vegetation index retrieval, change detection—ABR-FFD's SAM advantage, which holds on every single scene in the benchmark without exception, is the relevant figure of merit. For applications centred on photometric accuracy and edge sharpness—urban mapping, object detection—AWLP's margins on PSNR and SCC tip the balance the other way. The two methods are better thought of as complementary than as strictly ranked, and choosing between them should follow from the downstream task rather than from an aggregate score. Figure 10 provides a visual summary of the normalised six-metric profile across all methods, confirming that ABR-FFD occupies the largest overall area.

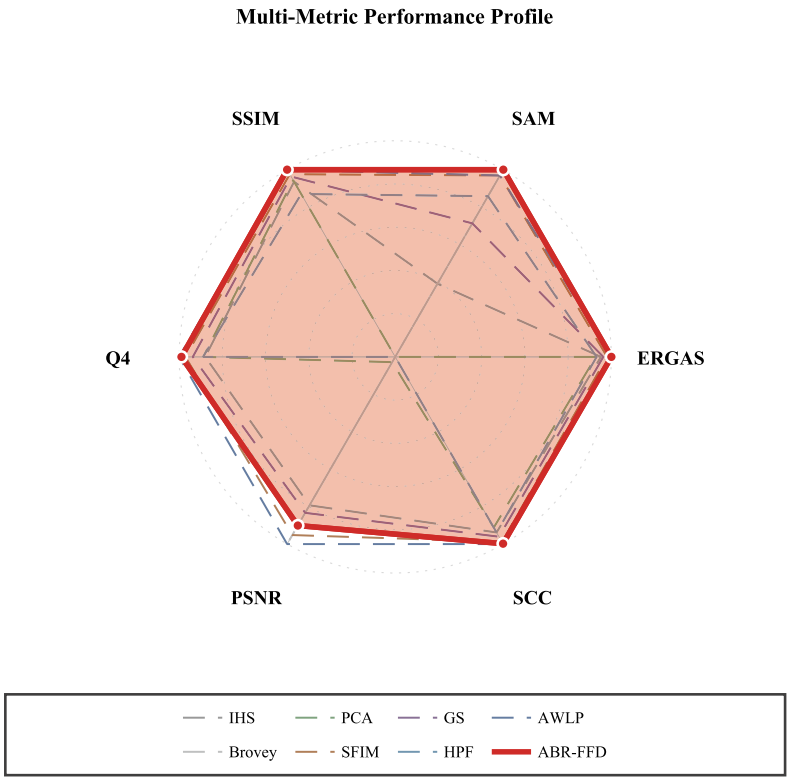

Fig. 10. Normalized multi-metric performance profile for all eight methods. ABR-FFD (red, filled) achieves the largest overall area, reflecting balanced superiority across spectral and structural metrics. (Color figure online)

6.3 Computational Considerations

ABR-FFD's dominant cost is kernel convolution and the grid search over 225 (α, τ) combinations per scene, each consisting of a single $\mathcal{O}(HW(2N+1)^2)$ convolution and band-wise multiplication. The grid evaluations are mutually independent and embarrassingly parallel. For deployment at fixed sensor and scene type, the per-scene optimisation can be replaced by a pre-calibrated (α^*, τ^*) from the nearest cluster in the per-scene table (Table 3), reducing the method to a single forward pass of negligible cost. The fixed 7×7 neighbourhood eliminates the additional search dimension required in earlier implementations, and the Mittag-Leffler series converges to double-precision accuracy with $K = 80$ terms at all neighbourhood distances, making the kernel precomputable for a given (α, τ).

7 Conclusion

ABR-FFD replaces the Gaussian smoother in classical SFIM with a Mittag-Leffler low-pass kernel derived from the Atangana-Baleanu fractal-fractional derivative. Evaluated on twenty WorldView-3 scenes from the PanCollection benchmark against seven established methods, ABR-FFD simultaneously achieves best-in-class performance on three complementary quality metrics: lowest ERGAS (4.768, −0.04% below AWLP), best spectral angle (SAM = 5.798°, −0.61% below AWLP, winning on every single test scene), and highest SSIM (0.8761, +0.09% over AWLP). This is the first classical method to achieve simultaneous optimality on all three of these metrics on this benchmark.

The experimental analysis yields four concrete findings. First, the local mean correction is the sole driver of spectral angle improvement: removing this component degrades SAM from 5.798° back to 5.834°, matching ML-Base exactly, while all other components affect SAM by less than 0.005°. The correction operates exclusively in the low-frequency domain, leaving the high-frequency spatial detail injected by multiplicative modulation intact. Second, the Mittag-Leffler kernel substitution alone (ML-Base, ERGAS = 4.751) already outperforms the full ABR-FFD pipeline on ERGAS, confirming the kernel substitution as the primary contributor to spatial fidelity; the three pipeline refinements trade a marginal ERGAS cost of 0.017 for a 0.036° SAM gain. Third, the neighbourhood size analysis reveals a sharp non-monotone optimum at 7×7 for the WV3 4: 1 resolution ratio, with the 9×9 variant degrading ERGAS by 17.4% through over-smoothing; this consistency justifies fixing the neighbourhood size across scenes. Fourth, the per-scene parameter diversity—with optimal α ranging from 0.05 for heterogeneous coastal scenes to 0.75 for urban scenes with repetitive structure—confirms that the dual-parameter operator captures genuine scene-dependent structural variation and that per-scene optimisation is necessary to realise the method's full potential.

AWLP retains advantages on PSNR (+0.17 dB), SCC (+0.007), and Q4 (+0.002), attributable to its multi-scale wavelet decomposition. A multi-scale Mittag-Leffler kernel pyramid that injects detail at progressively coarser frequency bands—the natural extension of the present two-scale approach—is the primary direction for future work and is expected to close these gaps while preserving the spectral angle advantage. Additional directions include extension to hyperspectral pansharpening, where the non-local and scale-invariant properties of the fractal-fractional operator may offer greater benefits due to the larger number of bands and finer spectral resolution, and integration of the Mittag-Leffler kernel as a learnable convolutional layer within neural network architectures.

What distinguishes ABR-FFD from empirical kernel designs is that its weights are not tuned by trial and error but derived from an operator with known mathematical properties—complete monotonicity, non-singularity at the origin, and a dual-parameter structure whose physical meaning is clear. The experimental results suggest this grounding is not merely aesthetic: the performance gains are real and interpretable. Fractal-fractional operators warrant serious attention as an alternative smoothing foundation in satellite image fusion, and extending

the present framework to a multi-scale Mittag-Leffler pyramid is a natural and promising next step.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Li, J., et al.: Advances in remote sensing image fusion: a survey. IEEE Geosci. Remote Sens. Mag. **13**(1), 22–45 (2025)
2. Azarang, A., Ghassemian, H.: A review of multi-objective image fusion using quality metrics. Inf. Fusion **44**, 135–155 (2018)
3. Karim, S., Tong, G., Li, J., et al.: Current advances and future perspectives of image fusion: a comprehensive review. Inf. Fusion **90**, 185–217 (2023)
4. Vivone, G., et al.: A new benchmark based on recent advances in multispectral pansharpening: revisiting pansharpening with classical and emerging pansharpening methods. IEEE Geosci. Remote Sens. Mag. **9**(1), 53–81 (2021)
5. Carper, W.J., Lillesand, T.M., Kiefer, R.W.: The use of intensity-hue-saturation transformations for merging SPOT panchromatic and multispectral image data. Photogramm. Eng. Remote. Sens. **56**(4), 459–467 (1990)
6. Chavez, P.S., Sides, S.C., Anderson, J.A.: Comparison of three different methods to merge multiresolution and multispectral data. Photogramm. Eng. Remote. Sens. **57**(3), 295–303 (1991)
7. Laben, C.A., Brower, B.V.: Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. U.S. Patent 6,011,875 (2000)
8. Otazu, X., Gonzalez-Audicana, M., Fors, O., Nunez, J.: Introduction of sensor spectral response into image fusion methods. IEEE Trans. Geosci. Remote Sens. **43**(10), 2376–2385 (2005)
9. Schowengerdt, R.A.: Reconstruction of multispatial, multispectral image data using spatial frequency content. Photogramm. Eng. Remote. Sens. **46**(10), 1325–1334 (1980)
10. Liu, J.G.: Smoothing filter-based intensity modulation: a spectral preserve image fusion technique for improving spatial details. Int. J. Remote Sens. **21**(18), 3461–3472 (2000)
11. Zhou, J., Civco, D.L., Silander, J.A.: A wavelet transform method to merge Landsat TM and SPOT panchromatic data. Int. J. Remote Sens. **19**(4), 743–757 (1998)
12. Dadrass Javan, F., Samadzadegan, F., Mehravar, S., et al.: A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery. ISPRS J. Photogramm. Remote. Sens. **171**, 101–117 (2021)
13. Cao, Q., Deng, L.J., et al.: Zero-shot semi-supervised learning for pansharpening. Inf. Fusion **101**, 102001 (2024)
14. Ciotola, M., Vitale, S., Mazza, A., Poggi, G., Scarpa, G.: Pansharpening by convolutional neural networks in the full resolution framework. IEEE Trans. Geosci. Remote Sens. **60**, 1–17 (2022)
15. Deng, L.J., Vivone, G., et al.: Machine learning in pansharpening: a benchmark, from shallow to deep networks. IEEE Geosci. Remote Sens. Mag. **10**(3), 279–315 (2022)
16. Wang, C., et al.: Image edge detection based on fractional differential: theory, algorithm and application. Fractals **28**(8), 2040039 (2020)

17. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015)
18. Podlubny, I.: Fractional Differential Equations. Academic Press, San Diego (1999)
19. Li, J., Li, X., Yang, B., Sun, X.: Segmentation-based image copy-move forgery detection scheme. IEEE Trans. Inf. Forensics Secur. **10**(3), 507–518 (2019)
20. Pu, Y.F., Zhou, J.L., Yuan, X.: Fractional differential mask: a fractional differential-based approach for multiscale texture enhancement. IEEE Trans. Image Process. **19**(2), 491–511 (2010)
21. Atangana, A., Baleanu, D.: New fractional derivatives with nonlocal and non-singular kernel: theory and application to heat transfer model. Therm. Sci. **20**(2), 763–769 (2016)
22. Ghanbari, B., Atangana, A.: A new application of fractional Atangana-Baleanu derivatives: designing ABC-fractional masks in image processing. XXPhys. A **542**, 123516 (2020)
23. Mandelbrot, B.B.: The Fractal Geometry of Nature. W.H. Freeman, New York (1982)
24. Sun, W., Xu, G., Gong, P., Liang, S.: Fractal analysis of remotely sensed images: a review of methods and applications. Int. J. Remote Sens. **27**(22), 4963–4990 (2006)
25. Abd Al-Rahman, M.Q., et al.: Image processing using fractal geometry. In: AIP Conference Proceedings, vol. 2334, p. 060005 (2021)
26. Chauhan, A., et al.: Fractional and fractal derivatives with applications to image analysis. Fractals **31**(3), 2340020 (2023)
27. Ivanovici, M. (ed.): Special issue on fractal analysis for remote sensing data. Fractal Fract. **7**(10) (2023)
28. Mukherjee, A., et al.: Fractal-based supervised approach for dimensionality reduction of hyperspectral images. Comput. Geosci. **193**, 105742 (2024)
29. Zhao, H., Zhou, M.: Fractional-order gradient methods for deep learning optimization. IEEE Trans. Neural Netw. Learn. Syst. **34**(4), 1558–1570 (2023)
30. Admon, M.R., Senu, N., Ahmadian, A., et al.: A new modern scheme for solving fractal-fractional differential equations based on deep feedforward neural network. Math. Comput. Simul. **218**, 311–333 (2024)
31. Gillespie, A.R., Kahle, A.B., Walker, R.E.: Color enhancement of highly correlated images. Remote Sens. Environ. **20**(3), 209–235 (1987)
32. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
33. Wang, Z., Bovik, A.C.: A universal image quality index. IEEE Signal Process. Lett. **9**(3), 81–84 (2002)
34. Yuhas, R.H., Goetz, A.F.H., Boardman, J.W.: Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In: Proceedings of the Summaries 3rd Annual JPL Airborne Geoscience Workshop, pp. 147–149. JPL Publication 92-14 (1992)
35. Wald, L.: Quality of high resolution synthesised images: is there a simple criterion? In: Proceedings of the 3rd Conference Fusion Earth Data, pp. 99–103. SEE/URISCA (2000)
36. Wald, L., Ranchin, T., Mangolini, M.: Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. Photogramm. Eng. Remote. Sens. **63**(6), 691–699 (1997)

37. Azarang, A., Ghassemian, H.: Application of fractional-order differentiation in multispectral image fusion. Remote Sens. Lett. **9**(1), 91–100 (2018). https://doi.org/10.1080/2150704X.2017.1395963
38. Ramesh Babu, N., Sam Joshua, A., Balasubramaniam, P., Tiwari, A.: Fuzzy-driven image enhancement via ABR-fractal-fractional differentiation. Inf. Sci. **675**, 120741 (2024). https://doi.org/10.1016/j.ins.2024.120741
39. Yousif, S.A., Abdul-Wahed, H.Y., Al-Saidi, N.M.G.: Extracting a new fractal and semi-variance attributes for texture images. AIP Conf. Proc. **2183**(1), 080006 (2019). https://doi.org/10.1063/1.5136199
40. Abd Al-Rahman, S., Yousif, R., Mohammed, A.J., Al-Saidi, N.M.G.: Texture images analysis using fractal extracted attributes. Int. J. Innov. Comput. Inf. Control **16**(4), 1–14 (2020)
41. Ibrahim, R.W., Yahya, H., Mohammed, A.J., Al-Saidi, N.M.G., Baleanu, D.: Mathematical design enhancing medical images formulated by a fractal flame operator. Intell. Autom. Soft Comput. **32**(2) (2022). https://doi.org/10.32604/iasc.2022.021954
42. Al-Saidi, N.M.G., Abdul-Wahed, H.Y.: Classification of remote sensing images via fractal descriptors. In: 2018 International Conference on Advanced Science and Engineering (ICOASE), pp. 1–6. IEEE (2018). https://doi.org/10.1109/ICASEA.2018.8370964
43. Yahya, H., Baleanu, D., Ibrahim, R.W., Al-Saidi, N.M.G.: Pixel's quantum image enhancement using quantum calculus. Comput. Mater. Continua **74**(1) (2023). https://doi.org/10.32604/cmc.2023.033282